

CUTOFF FOR RANDOM WALKS ON GRAPHS WITH BOTTLENECKS.

IOANNIS PAPAGEORGIOU

ABSTRACT. We examine the mixing time for random walks on graphs. In particular we are interested on investigating graphs with bottlenecks. Furthermore, the cutoff phenomenon is examined.

1. INTRODUCTION.

Assume X is an irreducible aperiodic Markov chain on some finite state space. Consider P to be the transition matrix and π the stationary distribution. At first define the total variation distance between two measures μ and ν to be

$$\|\mu - \nu\| = \sup_A |\mu(A) - \nu(A)|$$

Then for every $\epsilon > 0$ we define the ϵ -total variation mixing time as

$$t_{mix}(\epsilon) = \min\{t \geq 0 : \max_x \|P^t(x, \cdot) - \pi\| \leq \epsilon\}$$

The purpose of the current paper is to calculate upper and lower bounds for the mixing time of an irreducible Markov chain on graphs with bottlenecks. In particular, we are interested in determining conditions under which the upper and lower bounds of the mixing time are asymptotically equal. Furthermore, some outcomes in relation to the cutoff phenomena are obtained. We say that the sequence of events X^n exhibits total variation cutoff if for every $0 < \epsilon < 1$

$$(1.1) \quad \lim_{n \rightarrow \infty} \frac{t_{mix}^n(\epsilon)}{t_{mix}^n(1 - \epsilon)} = 1$$

In [P-S] the first example of a tree was constructed which exhibits total variation cutoff. The construction of the tree was based on placing a binary tree consisting of $N = n_k^3$ vertices at the origin of a line of n_k points, where $n_k = 2^{2^k}$. Then for every $j \in \{[k/2], \dots, k\}$ a binary tree T_j consisting of N/n_j vertices was placed at distance n_j from the origin. The purpose of the current paper is to generalise this result. We will consider two cases, one general case referred to as Case A

2000 *Mathematics Subject Classification.* 60J10.

Key words and phrases. cutoff, mixing time.

Address: Departamento de Matematica, Universidad de Buenos Aires, Pabellon II, Ciudad Universitaria Buenos Aires, Argentina.

Email: ipapageorgiou@dm.uba.ar, papyannis@yahoo.com.

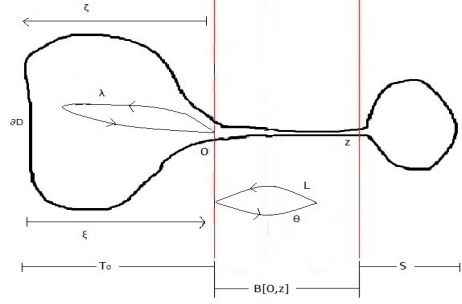


FIGURE 1. Case A.

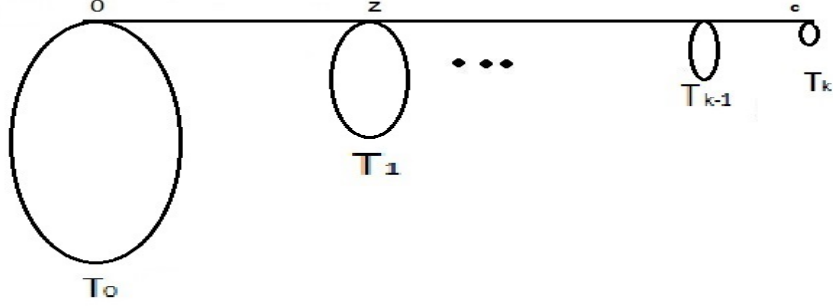


FIGURE 2. Case B.

(see figure 1) where we will consider two graphs connected with a bottleneck and another one referred to as Case B (see figure 2), where we will substitute the trees T_j in [P-S] by finite graphs T_j in such a way that a bottleneck is observed between the T_j 's. We will denote by $B[0, z]$ the bottleneck of the graph, the consistent part of the graph between 0 and z . $T_i, S_i, B[0, z]$ all increase in size as k increases. We will denote $D = T_0 \cup B[0, z]$ while the complement of D will be denoted as $S = D^c$. Furthermore, we will denote ∂D to be the edge of D , i.e the points of T_0 that are further from 0. In Case A we will denote c the edge of S , while in case B c will be the connecting point between T_k and the rest of the graph. We will consider the size of the bottleneck and S to be relatively small compared with how the size of T_0 increases, in such a way that

$$(C1) \quad \pi(B[0, z] \cup S) \rightarrow 0 \text{ as } k \rightarrow \infty$$

We define $1 - p$ to be the probability at every return to T_0 the random walk to hit the edge ∂D of T_0 before exiting T_0 . For instance, if T_0 is a binary tree, as is

the case in [P-S], $1 - p \geq \frac{1}{3}$. For any $x \in \Omega$, by τ_x we define the first time we reach x

$$\tau_x = \inf\{t : x_t = x\}$$

Thus, τ_z denotes the first time we reach the boundary between S and D . Furthermore, we define

$$\tau_x^* = \inf\{s \geq \tau_0 : x_s = x\}$$

We will now look at the main conditions and results of the paper. If for some γ increasing on k the following condition holds

$$(H1) \quad \mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{\text{Var}_c(\tau_0)} \geq \mathbb{E}_{\partial D}(\tau_0) + \delta \sqrt{\text{Var}_{\partial D}(\tau_0)}$$

then we show that $t_{mix}(\epsilon) \sim \mathbb{E}_c(\tau_0)$. This is the results presented in Theorem 1.1. In this category belongs the example presented in [P-S], as explained in section 4.1. Then we look on graphs that (H1) does not hold. It appears that with some additional conditions the statement of the theorem derived under (H1) still holds true. Assume that the opposite inequality of (H1) holds, that is that

$$\mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{\text{Var}_c(\tau_0)} \leq \mathbb{E}_{\partial D}(\tau_0) + \delta \sqrt{\text{Var}_{\partial D}(\tau_0)}$$

We will determine conditions so that an asymptotic estimate of the mixing time can be obtained in this case as well. Denote $t_S = \mathbb{E}_c(\tau_0) - \gamma \sqrt{\text{Var}_c(\tau_0)}$. Assume that

$$(1.2) \quad \Phi = \sum_{x \in S} \frac{\pi(x)}{\pi(S)} P_x(t_S > \tau_0) < 1$$

and in addition D and S are such that the following inequality holds (H2):

$$\frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{\text{Var}_c(\tau_0)}}{A} \geq \mathbb{E}_{\partial D}(\tau_0) + \delta \sqrt{\text{Var}_{\partial D}(\tau_0)} \geq \mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{\text{Var}_c(\tau_0)}$$

where $A = \frac{\Phi}{1 - \pi(S) - \epsilon} < 1$ for every $\epsilon < 1 - \pi(S) - \Phi$. Then $t_{mix}(\epsilon) = \frac{\mathbb{E}_c(\tau_0)}{A}$. This will be the subject of Theorem 1.3. One should notice that the right inequality of (H2) is nothing else than the inverse of (H1). Furthermore, A do not need to be a constant but as shown on the two examples presented in sections 4.2 and 4.3 A can be a decreasing function such that $A^{-1} \rightarrow \infty$ as $k \rightarrow \infty$. In this way we can even consider examples that go far from (H1) such that

$$\frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{\text{Var}_c(\tau_0)}}{\mathbb{E}_{\partial D}(\tau_0) + \delta \sqrt{\text{Var}_{\partial D}(\tau_0)}} \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

For both cases investigated in theorems 1.1 and 1.3 cutoff properties are proven in corollaries 1.2 and 1.4.

Some further assumptions on the graph. In relation to the bottleneck we also define

$$(1.3) \quad L = \sum_{s=\tau_0}^{\tau_z^*} \mathcal{I}(x_{s-1} \notin T_0, x_s \in T_0)$$

the number of returns to T_0 in the time interval $[\tau_0, \tau_z^*]$. We also assume that the part of it between 0 and z is such that there exists an increasing function h such that

$$(C2) \quad P(L \leq z) \leq \frac{1}{h(z)}$$

For the analogue of the example in [P-S] where the bottle neck is a line, $h(a) = a$. Through out this paper we will use the symbol \prec to denote stochastic domination, i.e. we will write

$$A \prec B \Leftrightarrow \lim_{k \rightarrow +\infty} \frac{A}{B} = 0 \quad \text{a.e.}$$

Furthermore, we will write

$$A \asymp B \Leftrightarrow \exists \text{ constant } c \text{ s.t. } A = cB + o(1)$$

and

$$A \lesssim B \Leftrightarrow \exists \text{ constant } c \text{ s.t. } A \leq cB + o(1)$$

We define $(\theta_i)_i$ to be iid variables distributed as the length of a random walk on $B[0, z]$ starting from 0 conditioning not to hit z and $(\lambda_{i,j}, j)$ to be iid random variables distributed as the length of a random walk from 0 on T_0 conditioning not to hit ∂T_0 the edge of T_0 . We will also assume

$$(C3) \quad \begin{aligned} \mathbb{E}_0(\tau_{\partial D}) &\leq \mathbb{E}_{\partial D}(\tau_0) \\ \mathbb{E}_{\partial T_i}(n_i) &\leq \mathbb{E}_{\partial D}(\tau_0) \quad \forall i \end{aligned}$$

Conditions (C3) are reasonable enough since they roughly state that the time needed to traverse the big set of vertices with the complex structure T_0 is bigger than the smaller T_i 's, as well as that when in T_0 the random walk moves with bigger probability towards the edge than the bottleneck. The main theorem about the mixing time follows.

Theorem 1.1. *Assume conditions (C) and*

$$(H)(a) \quad \gamma \sqrt{\text{Var}_c(\tau_0)} \prec \mathbb{E}_c(\tau_0) \quad \text{and} \quad (b) \quad \mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L] \mathbb{E}[\lambda_{i,j} + \theta_i] \prec \gamma \sqrt{\text{Var}_c(\tau_0)}$$

If γ and δ are increasing functions such that

$$(H1) \quad \mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{\text{Var}_c(\tau_0)} \geq \mathbb{E}_{\partial D}(\tau_0) + \delta \sqrt{\text{Var}_{\partial D}(\tau_0)}$$

then for every $0 < \epsilon < 1$ the mixing time for the random walk on the graph is

$$t_{\text{mix}}(\epsilon) \sim \mathbb{E}_c(\tau_0)$$

In relation then to the cutoff phenomenon we obtain the following corollary.

Corollary 1.2. *Under the conditions of Theorem 1.1 the random walk exhibits variation cutoff, that is for every $0 < \epsilon < 1$*

$$\lim_{n \rightarrow \infty} \frac{t_{mix}^n(\epsilon)}{t_{mix}^n(1 - \epsilon)} = 1$$

For an example that satisfies the conditions of Theorem 1.1 one can look on graphs similar to the one presented in [P-S] as described in section 4.1.

In the case where condition (H1) is not satisfied then an analogue result still holds as presented on the next theorem.

Theorem 1.3. *Assume that $\Phi = \sum_{x \in S} \frac{\pi(x)}{\pi(S)} P_x(t_S > \tau_0) < 1$. Assume conditions (C) and*

$$(H) \text{ (a) } \gamma \sqrt{Var_c(\tau_0)} \prec \mathbb{E}_c(\tau_0) \text{ and (b) } \mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L] \mathbb{E}[\lambda_{i,j} + \theta_i] \prec \frac{\gamma \sqrt{Var_c(\tau_0)}}{A}$$

If γ and δ are increasing functions such that

$$(H2) \quad \frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{Var_c(\tau_0)}}{A} \geq \mathbb{E}_{\partial D}(\tau_0) + \delta \sqrt{Var_{\partial D}(\tau_0)} \geq \mathbb{E}_c(\tau_0) + \frac{\gamma}{2} \sqrt{Var_c(\tau_0)}$$

then for every $\epsilon < 1 - \pi(S) - \Phi$

$$t_{mix}(\epsilon) = \frac{\mathbb{E}_c(\tau_0)}{A}$$

where $A = \frac{\Phi}{1 - \pi(S) - \epsilon} < 1$.

In sections 4.2 and 4.3 two examples that satisfy the conditions of Theorem 1.3 are presented. As an outcome of the theorem, we obtain the following corollary about variational cutoff.

Corollary 1.4. *Under the conditions of Theorem 1.3 the random walk exhibits variation cutoff, that is for every $0 < \epsilon < 1 - \pi(S) - \Phi$*

$$\lim_{n \rightarrow \infty} \frac{t_{mix}^n(\epsilon)}{t_{mix}^n(1 - \epsilon)} = 1$$

It should be noted that the variational cutoff phenomenon presented in Corollary 1.4 is weaker than the one presented in [P-S] and Corollary 1.2 since in the first two $\epsilon \in (0, 1)$ while in the last corollary $\epsilon \in (0, 1 - \pi(S) - \Phi)$. However, if one constructs a graph such that $\Phi \rightarrow 0$ as $k \rightarrow \infty$, since $\pi(S) \rightarrow \infty$, we obtain the cutoff asymptotically for $\epsilon \in (0, 1)$. For examples of graphs with $\Phi \rightarrow 0$ one can look on paradigm 2 and 3 of section 4.

In the following theorem we present a weaker version of Theorem 1.3, which will be used to show the examples of section 4.

Theorem 1.5. *Assume that $\Phi = t_S \sum_{x \in S} \sum_{y \notin S} \mu_S(x) P(x, y) < 1$ and conditions (C) and (H). If*

$$\mathbb{E}_{\partial D}(\tau_0) \prec \delta \sqrt{\text{Var}_{\partial D}(\tau_0)}$$

and

$$\mathbb{E}_c(\tau_0) \leq \delta \sqrt{\text{Var}_{\partial D}(\tau_0)} \leq \frac{\mathbb{E}_c(\tau_0)}{A}$$

then for every $\epsilon < 1 - \pi(S) - \Phi$

$$t_{mix}(\epsilon) = \frac{\mathbb{E}_c(\tau_0)}{A}$$

where $A = \frac{\Phi}{1 - \pi(S) - \epsilon} < 1$.

A few words about the proof of the two theorems and the structure of the paper. In order to find the asymptotic limits of the mixing time t_{mix} we will calculate upper and lower bounds for $t_{mix}(\epsilon)$. Under the conditions of Theorem 1.1, for $0 < \epsilon < 1$ we will show that

$$(1.4) \quad \mathbb{E}_c(\tau_0) - \gamma \sqrt{\text{Var}_c(\tau_0)} \leq t_{mix}(\epsilon) \leq \mathbb{E}_c(\tau_0) + \gamma \sqrt{\text{Var}_c(\tau_0)}$$

Then the mixing time and the cutoff property follow because of (H)(a).

Under the conditions of Theorem 1.3, we will show that for every $\epsilon \in (0, 1 - \pi(S) - \Phi)$

$$(1.5) \quad \frac{\mathbb{E}_c(\tau_0) - \gamma \sqrt{\text{Var}_c(\tau_0)}}{A} \leq t_{mix}(\epsilon) \leq \frac{\mathbb{E}_c(\tau_0) + \gamma \sqrt{\text{Var}_c(\tau_0)}}{A}$$

Then the mixing time and the cutoff property follow again because of (H)(a).

The lower bound of both (1.4) and (1.5) will be shown in Lemma 2.1 and Proposition 2.2 of section 2 respectively. The upper bound of (1.4) and (1.5) follows from Proposition 3.2 for $A = 1$ and $A = \frac{\Phi}{1 - \pi(S) - \epsilon} < 1$ respectively.

2. LOWER BOUNDS.

In this section we present two lower bounds for the mixing time t_{mix} . The first lower bound is presented in Lemma 2.1. Then under the condition $\Phi = \sum_{x \in S} \frac{\pi(x)}{\pi(S)} P_x(t_S > \tau_0) < 1$ we prove in Proposition 2.2 a sharper lower bound. The first lower bound for the mixing time follows.

Lemma 2.1. *The following lower bound for the mixing time holds:*

$$t_{mix}(\epsilon) \geq \mathbb{E}_c(\tau_0) - \gamma \sqrt{\text{Var}_c(\tau_0)}$$

The proof of Lemma 2.1 is presented in [P-S]. We will use this bound in order to show a sharper lower bound on the following proposition.

Denote π_S the restriction of π to S , $\pi_S(A) = \pi(S \cap A)$ and $\mu_S(A) = \frac{\pi_S(A)}{\pi(S)}$. In order to prove the second sharper lower bound we will use the approach of [L-P-W] for graphs with one bottleneck.

The main result related to the lower bound of the mixing time follows in the next proposition.

Proposition 2.2. *Assume that $\Phi = t_S \sum_{x \in S} \sum_{y \notin S} \mu_S(x) P(x, y) < 1$. If $S \subset \Omega$ such that $\pi(S) < 1$ then for every $\epsilon < 1 - \pi(S) - \Phi$*

$$t_{mix}(\epsilon) \geq \frac{\mathbb{E}_c(\tau_0) - \gamma \sqrt{Var_c(\tau_0)}}{A}$$

where $A = \frac{\Phi}{1 - \pi(S) - \epsilon}$.

Proof. Since from Lemma 2.1 we know that $t_{mix} \geq \mathbb{E}_c(\tau_0) - \gamma \sqrt{Var_c(\tau_0)} = t_S$, there exists an $m \geq 1$ such that $t_{mix(\epsilon)} = mt_S$. In order to bound the total variational distance $\| \mu_S P^{mt_S} - \mu_S \|_{TV}$ we can use the following bound

$$\| \mu_S P^t - \mu_S \|_{TV} \leq t \Phi(S), \quad t \in \mathbb{N}$$

where $\Phi(S) = \sum_{x \in S} \sum_{y \notin S} \mu_S(x) P(x, y)$, (see (7.14)-(7.15) from [L-P-W]). We then obtain

$$(2.1) \quad \| \mu_S P^{t_{mix}} - \mu_S \|_{TV} \leq t_{mix} \Phi(S) = mt_S \Phi(S) = m \Phi$$

where we have denoted $\Phi = t_S \Phi(S)$. We have

$$\| \mu_S - \pi \|_{TV} = \max_{A \subset \Omega} | \mu_S(A) - \pi(A) | \geq | \mu_S(S^c) - \pi(S^c) | = \pi(S^c)$$

since $\mu_S(S^c) = 0$. So, we can write

$$1 - \pi(S) \leq \| \mu_S - \pi \|_{TV} \leq \| \mu_S P^{t_{mix}} - \mu_S \|_{TV} + \| \mu_S P^{t_{mix}} - \pi \|_{TV}$$

which leads to

$$1 - \pi(S) \leq \| \mu_S P^{t_{mix}} - \mu_S \|_{TV} + \epsilon$$

If we bound the first term on the right hand by (2.1) we then have

$$(2.2) \quad 1 - \pi(S) \leq m \Phi + \epsilon$$

which after substituting $m = \frac{t_{mix(\epsilon)}}{t_S}$ gives the following lower bound for $t_{mix(\epsilon)}$

$$t_{mix(\epsilon)} \geq \frac{t_S}{\frac{\Phi}{1 - \pi(S) - \epsilon}}$$

□

3. UPPER BOUND FOR THE CUTOFF CASE.

In this section we calculate the upper bound for the mixing time. Technics from [P-S] will be closely followed. We start with a technical lemma.

Lemma 3.1. *Assume (C2), (H2) and (H)(b). If we denote $t' = \frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{2}\sqrt{\text{Var}_c(\tau_0)}}{A}$, and define the set $B = \{\tau_0 < t'\}$. Then*

$$(3.1) \quad P_x(B^c) \leq \frac{16}{\gamma^2} + \frac{1}{\delta^2}$$

Proof. We will first consider $x \in S \cup B[0, z]$. To show (3.1) we will distinguish on the two different cases of graphs denoted as Case A and B, shown on figures 1 and 2 respectively.

In the Case A we have

$$(3.2) \quad \begin{aligned} P_x(B^c) &= P_x(\tau_0 \geq t') \leq P_c(\tau_0 \geq t') = P_c(\tau_0 \geq \frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{2}\sqrt{\text{Var}_c(\tau_0)}}{A}) \\ &\leq P_c(\tau_0 \geq \mathbb{E}_c(\tau_0) + \frac{\gamma}{2}\sqrt{\text{Var}_c(\tau_0)}) \leq \frac{4}{\gamma^2} \end{aligned}$$

where above we used first that $A \leq 1$ and then Chebyshev's inequality.

We will show the same for every $x \in S \cup B[0, z]$ for graphs in Case B. Define τ_1 the time it gets to hit $[0, c]$ and τ_2 the time it takes to hit 0 starting from X_{τ_1} . Then

$$(3.3) \quad \begin{aligned} P_x(B^c) &= P_x(\tau_0 \geq t') = P_x(\tau_0 \geq \frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{2}\sqrt{\text{Var}_c(\tau_0)}}{A}) \leq \\ &\leq P_x(\tau_1 \geq \frac{\gamma}{4} \frac{\sqrt{\text{Var}_c(\tau_0)}}{A}) + P_x(\tau_2 \geq \frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{4}\sqrt{\text{Var}_c(\tau_0)}}{A}) \end{aligned}$$

For the first term on the right hand side of (3.3) we can use Markov inequality to get

$$P_x(\tau_1 \geq \frac{\gamma}{4} \frac{\sqrt{\text{Var}_c(\tau_0)}}{A}) \leq \frac{4\mathbb{E}_x(\tau_1)}{\frac{\gamma\sqrt{\text{Var}_c(\tau_0)}}{A}}$$

But for any $x \in T_i$

$$\mathbb{E}_x(\tau_1) \leq \mathbb{E}_{\partial T_i}(n_i) \leq \mathbb{E}_{\partial D}(\tau_0)$$

because of (C3). This leads to

$$(3.4) \quad P_x(\tau_1 \geq \frac{\gamma}{4} \frac{\sqrt{\text{Var}_c(\tau_0)}}{A}) \leq \frac{4\mathbb{E}_{\partial D}(\tau_0)}{\frac{\gamma\sqrt{\text{Var}_c(\tau_0)}}{A}} = 0(1)$$

because of (H)(b).

For the second term on the right hand side of (3.3) we have

$$\begin{aligned}
 P_x(\tau_2 \geq \frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{4}\sqrt{Var_c(\tau_0)}}{A}) &\leq P_x(\tau_2 \geq \mathbb{E}_c(\tau_0) + \frac{\gamma}{4}\sqrt{Var_c(\tau_0)}) \\
 &\leq P_c(\tau_0 \geq \mathbb{E}_c(\tau_0) + \frac{\gamma}{4}\sqrt{Var_c(\tau_0)}) \\
 &\leq \frac{16}{\gamma^2}
 \end{aligned}
 \tag{3.5}$$

where above we used that $A < 1$ and applied Chebyshevs inequality. Finally, putting (3.4) and (3.5) in (3.3) we obtain that for graphs in Case B, for every $x \in S \cup B[0, z]$

$$P_x(B^c) \lesssim \frac{16}{\gamma^2} \tag{3.6}$$

From (3.2) and (3.6) we obtain that in both Cases A and B, for every $x \in S \cup B[0, z]$ one has

$$P_x(B^c) \lesssim \frac{16}{\gamma^2} \tag{3.7}$$

For both Cases A and B, when $x \in T_0$ we have

$$\begin{aligned}
 P_x(B^c) &= P_x(\tau_0 \geq t') \leq P_{\partial D}(\tau_0 \geq t') \leq \\
 &\leq P_{\partial D}(\tau_0 \geq \mathbb{E}_{\partial D}(\tau_0) + \delta \sqrt{Var_{\partial D}(\tau_0)}) \leq \frac{1}{\delta^2}
 \end{aligned}
 \tag{3.8}$$

where above we used (H2) and Chebyshev's inequality. From (3.7) and (3.8) we obtain (3.1) for every $x \in D \cup S$. \square

The main result about the upper bound of the mixing time follows.

Proposition 3.2. *Assume (H) and that*

$$(H2) \quad \frac{\mathbb{E}_c(\tau_0) + \frac{\gamma}{2}\sqrt{Var_c(\tau_0)}}{A} \geq \mathbb{E}_{\partial D}(\tau_0) + \delta\sqrt{Var_{\partial D}(\tau_0)}$$

for some $A \leq 1$. Then

$$t_{mix} \leq \frac{\mathbb{E}_c(\tau_0) + \gamma\sqrt{Var_c(\tau_0)}}{A}$$

Proof. Denote $t = \frac{\mathbb{E}_c(\tau_0) + \gamma\sqrt{Var_c(\tau_0)}}{A}$. We will consider the following coupling. Assume $X_0 = x$ and $Y_0 \sim \pi$. We let X and Y move independently until the first time X hits 0. Then they still continue both moving independently until the moment they collide or reach the same level at T_0 . In this case the coupling changes to the following. X keeps moving as an aperiodic random walk while Y moves closer or further from 0 if X moves closer or further respectively. Define τ to be the coupling time.

Then define E to be the event that after hitting 0 for the first time it reaches the edge ∂D of T_0 before hitting z , i.e.

$$E = \{\tau_{\partial D}^* < \tau_z^*\}$$

as well as the events

$$A_L = \{L > z\}$$

If we define

$$S = \inf\{s \geq \tau_{\partial D}^* : x_s = 0\}$$

then on the event E

$$S - \tau_0 \prec \sum_{i=1}^L \theta_i + \sum_{i=1}^L \sum_{j=1}^{G_i} \lambda_{i,j} + \zeta + \xi$$

where we recall $(\theta_i)_i$ are iid variables distributed as the length of a random walk on $B[0, z]$ starting from 0 conditioning not to hit z and $(\lambda_{i,j})$ are iid random variables distributed as the length of a random walk from 0 on T_0 conditioning not to hit ∂T_0 the edge of T_0 . $(G_i)_i$ is a random variable with probability of success $\frac{\#\{j: j \sim 0, j \in B[0, z]\}}{\#\{j: j \sim 0\}}$ and ξ is a random variable distributed as the commute time between the edge ∂T_0 and 0. If we use Wald's identity we obtain

$$(3.9) \quad \mathbb{E}[(S - \tau_0)\mathcal{I}_E] \leq \mathbb{E}[L]\mathbb{E}[\theta_i] + \mathbb{E}[L]\mathbb{E}[G_i]\mathbb{E}[\lambda_{i,j}] + \mathbb{E}[\zeta] + \mathbb{E}[\xi] \prec \\ \mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L]\mathbb{E}[\lambda_{i,j}] + \mathbb{E}[L]\mathbb{E}[\theta_i]$$

where above we used (C3) for $\mathbb{E}[\xi] = \mathbb{E}_{\partial D}(\tau_0)$ and $\mathbb{E}[\zeta] = \mathbb{E}_0(\tau_{\partial D})$.

We compute

$$P(\tau > t) \leq P(\tau > t, A_L) + \frac{1}{h(z)} \\ \leq P(\tau > t, A_L, E) + \frac{1}{h(z)} + p^{h(z)}$$

where in the first inequality we used (C2). Then we obtain

$$P(\tau > t) \leq P(\tau > t, A_L, E) + \frac{1}{h(z)} + p^{h(z)}$$

If we use Lemma 3.1 the last one can be bounded by

$$(3.10) \quad P(\tau > t) \leq P(\tau > t, A_L, E, B) + \frac{16}{\gamma^2} + \frac{1}{\delta^2} + \frac{1}{h(z)} + p^{h(z)}$$

Now define $F = \{S - \tau_0 > \frac{\gamma}{2} \sqrt{\frac{\text{Var}_c(\tau_0)}{A}}\}$ and $M = \{Y_{\tau_0} \in D\}$. Then $P(M^c) = o(1)$ as $k \rightarrow \infty$ since at time τ_0 the random walk Y is stationary, and so, because of (C1) we have that the stationary probability of T_0 is $1 - o(1)$. We observe that on the events E and M the two walks X and Y must have coalesced by time S . Therefore

$$B \cap F^c \subset \{\tau < t\}$$

This implies that

$$\begin{aligned} P(\tau > t, A_L, E, B) &\leq P(\tau > t, A_L, E, M, B) + 0(1) = P(\tau > t, A_L, E, M, B, F) \leq \\ &\leq P(E, F) = P(E, \{S - \tau_0 > \frac{\gamma}{2} \sqrt{Var_c(\tau_0)}\}) \end{aligned}$$

From Markov inequality

$$P(\tau > t, A_L, E, C, B) \leq \frac{2\mathbb{E}[(S - \tau_0)\mathcal{I}_E]}{\frac{\gamma\sqrt{Var_c(\tau_0)}}{A}}$$

If we now use (3.9) we obtain

$$(3.11) \quad P(\tau > t, A_L, E, C, B) \leq 2 \frac{\mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L]\mathbb{E}[\lambda_{i,j}] + \mathbb{E}[L]\mathbb{E}[\theta_i]}{\frac{\gamma\sqrt{Var_c(\tau_0)}}{A}}$$

Combining together (3.10) and (3.11) we get

$$P(\tau > t) \prec 2 \frac{\mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L]\mathbb{E}[\lambda_{i,j}] + \mathbb{E}[L]\mathbb{E}[\theta_i]}{\frac{\gamma\sqrt{Var_c(\tau_0)}}{A}}$$

If one takes under account that

$$\max_x \|P^t(x, \cdot) - \pi\| \leq P(\tau > t)$$

(see [L-P-W]) we eventually obtain

$$\max_x \|P^t(x, \cdot) - \pi\| \prec 2 \frac{\mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L]\mathbb{E}[\lambda_{i,j}] + \mathbb{E}[L]\mathbb{E}[\theta_i]}{\frac{\gamma\sqrt{Var_c(\tau_0)}}{A}}$$

which because of (H)(b) results to $t_{mix} \leq \frac{\mathbb{E}_c(\tau_0) + \gamma\sqrt{Var_c(\tau_0)}}{A}$. □

4. PARADIGMS

In this section we present examples for the two main theorems. At first in paradigm 1 an example that satisfies the conditions of Theorem 1.1 is presented. Then in sections 4.2 and 4.3, two examples that satisfy the conditions of Theorem 1.3 are presented. At the first one, paradigms 2, we consider a graph with

$$\mathbb{E}_{\partial D}(\tau_0) \leq \sqrt{Var_{\partial D}(\tau_0)}$$

while at the second one, paradigm 3, a graph with

$$\mathbb{E}_{\partial D}(\tau_0) > \sqrt{Var_{\partial D}(\tau_0)}$$

For both examples we establish $A^{-1} \rightarrow \infty$ as $k \rightarrow \infty$.

4.1. paradigm 1. The example presented in [P-S].

We will show that conditions (H1) is satisfied. Since in a binary tree one has

$$\sqrt{\text{Var}_{\partial D}(\tau_0)} \lesssim \mathbb{E}_{\partial D}(\tau_0)$$

it is sufficient to show that

$$(4.1) \quad \mathbb{E}_{\partial D}(\tau_0) \lesssim \sqrt{\text{Var}_c(\tau_0)}$$

Since N is the size of binary tree T_0 , one has that $N = \sum_{j=0}^l 2^j$, where l is the number of levels of T_0 , i.e. $l = \log(N - 1) - 1$. Concerning (4.1), one can think of the random walk from the leafs ∂T_0 of the binary tree to the origin 0 as a walk from 0 to l on the half line $[0, +\infty]$, with a reflective boundary at 0 and probabilities $\frac{2}{3}$ and $\frac{1}{3}$ towards the left and the right respectively at any other point of the line. Since, for this one dimensional random walk the hitting time satisfies

$$l^2 \lesssim \mathbb{E}_0(\tau_l) \lesssim l^3$$

we obtain for T_0 that

$$(4.2) \quad l^2 \asymp (\log(N - 1) - 1)^2 \lesssim \mathbb{E}_{\partial D}(\tau_0) \lesssim (\log(N - 1) - 1)^3 \asymp l^3$$

On the other hand we know that (see [P-S])

$$(4.3) \quad \sqrt{\text{Var}_c(\tau_0)} \asymp N\sqrt{k}$$

From (4.2) and (4.3) inequality (4.1) follows for appropriately large $N = n_k^3$ so that

$$(\log(N - 1) - 1)^3 \lesssim N\sqrt{k}$$

or if we substitute $N = n_k^3$

$$n_k^3 < e^{n_k k^{\frac{1}{6}} + 1}$$

which is true for n_k sufficiently large. The rest of the conditions are easily verified directly from [P-S].

4.2. paradigm 2. We will construct a graph based partly on the graph \mathcal{T} presented in [P-S]. Let $n_k = 2^{2^j}$ and consider the line $[0, n_k + 1]$. Then for all $j \in \{[\frac{k}{2}], \dots, k\}$ place a binary tree at distance n_j from the origin consisting of $\frac{N}{n_j}$ vertices. We denote this construction as G_0 . In this way, the part of G_0 contained between $[0, n_k]$, i.e. $G_0 \setminus \{n_k + 1\}$, is equal with $\mathcal{T} \setminus T_0$ from [P-S], where T_0 is the big binary tree at 0 of \mathcal{T} . Then consider r identical copies of G_0 and glue them together at 0 and $n_k + 1$ as shown on figure 3. We also consider q copies of a line $[-m, 0]$ and connect them with the previous construction at 0 and together at $-m$. In this way, we can consider T_0 to be the part of graph between $[-m, 0]$ with $\partial T_0 = -m$, the bottleneck $B[0, z]$ to be the part of the graph between 0 and $n_{[\frac{k}{2}]}$, while S is the part between $n_{[\frac{k}{2}]}$ and $n_k + 1$. In this way $c = n_k + 1$ and $z = n_{[\frac{n_k}{2}]}$.

We will determine m, r, q, γ, δ so that the conditions of Theorem 1.5 hold. But first, we will place conditions on the graph so that condition (H1) of Theorem 1.1

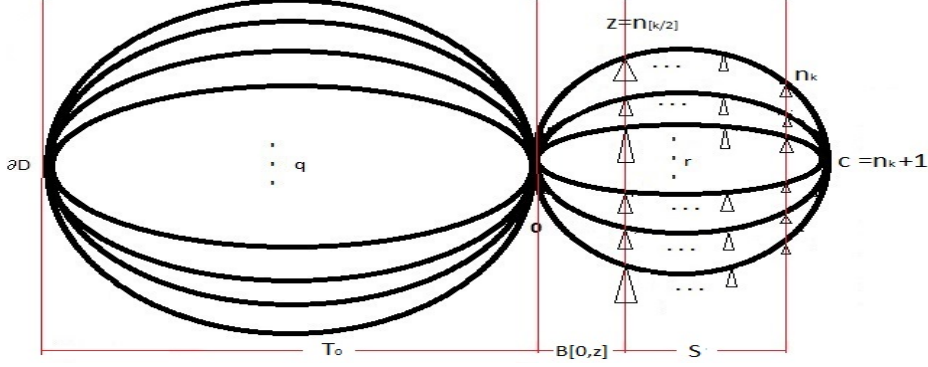


FIGURE 3. Paradigm 2

is not true. If X is a simple random walk on the interval $[-m, 0]$ starting from $-m$, then the mean and the variance of the time it gets to reach 0 from m for the first time are $\mathbb{E}_{-m}(\tau_0) = m^2$ and $Var_{-m}(\tau_0) = m^4$ respectively. Since D consists of q lines $[-m, 0]$ connected on 0 and $-m$, and when at $-m$ we move at any of the q branches $[-m, 0]$ with the same probability $\frac{1}{q}$ we obtain that

$$\mathbb{E}_{\partial D}(\tau_0) = m^2$$

while

$$Var_{\partial D}(\tau_0) = m^4$$

From [P-S] we know that

$$\sqrt{Var_c(\tau_0)} \asymp N\sqrt{k}$$

and

$$\mathbb{E}_c(\tau_0) = 6Nk + o(N\sqrt{k}) \asymp 6Nk$$

If one chooses

$$(4.4) \quad m^2 = 6Nk$$

then (H1) is not true since $\delta\sqrt{Var_{\partial D}(\tau_0)} > \mathbb{E}_c(\tau_0)$ for any $\delta > 1$. We also notice that for any $\delta = k^t$ with $t > 0$ one of the main conditions

$$\mathbb{E}_{\partial D}(\tau_0) \prec \delta\sqrt{Var_{\partial D}(\tau_0)}$$

is satisfied for k large enough. Furthermore, for any $\gamma = k^p$ with $p < \frac{1}{2}$ we get

$$(4.5) \quad \gamma\sqrt{Var_c(\tau_0)} \prec \mathbb{E}_c(\tau_0)$$

In addition, from (4.5) we get that

$$(4.6) \quad t_S = \mathbb{E}_c(\tau_0) - \gamma\sqrt{Var_c(\tau_0)} \prec 6Nk$$

At first we will show that for appropriate $r > 0$ one can obtain $A = \frac{\Phi}{1-\pi(S)-\epsilon} < \frac{1}{k^s}$ for some $s > 0$. Then $A \rightarrow \infty$ as $k \rightarrow \infty$.

If we choose $\Phi < 1$, then for every $\epsilon \leq 1-\pi(S)-\sqrt{\Phi}$ one has that $A = \frac{\Phi}{1-\pi(S)-\epsilon} \leq \sqrt{\Phi}$. Thus, it suffices to show that $\Phi = t_S \sum_{x \in S} \sum_{y \notin S} \mu_S(x) P(x, y) < \frac{1}{k^{2s}}$. For that we compute

$$(4.7) \quad \sum_{x \in S} \sum_{y \notin S} \mu_S(x) P(x, y) = \frac{q}{(q+r)|S|}$$

where the size of S is

$$(4.8) \quad |S| = r(n_k + 1 + N(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j})) \asymp rN(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j})$$

Combining together (4.6), (4.7) and (4.8) we get

$$\Phi \asymp \frac{6Nkq(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j})^{-1}}{(q+r)rN}$$

So, if we choose

$$(4.9) \quad r = 12k^{2s+1}(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j})^{-1}$$

we obtain

$$A \leq \sqrt{\Phi} \leq \frac{1}{k^s} < 1$$

for k sufficiently large. We will now determine parameters so that conditions (C) are satisfied.

For (C1) we compute

$$\pi(B[0, z] \cup S) = \frac{|S| + |B[0, z]|}{|S| + |B[0, z]| + |T_0|}$$

For this to vanish as $k \rightarrow \infty$ we need

$$(4.10) \quad |S| + |B[0, z]| \prec |T_0|$$

But $|S| + |B[0, z]| \asymp rN(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j})$ and $|T_0| \asymp q(m-2)$. For (C1) to hold true we need q large enough so that

$$\frac{rN}{m-2}(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j}) \prec q$$

If replace r by (4.9) and m by (4.4) we derive

$$(4.11) \quad \frac{12k^{2s+1}N}{\sqrt{6Nk}-2} \prec q$$

For (C2) we recall that

$$L = \sum_{s=\tau_0}^{\tau_z^*} \mathcal{I}(x_{s-1} \notin T_0, x_s \in T_0) = \sum_{s=\tau_0}^{\tau_{n_{[k/2]}^*}} \mathcal{I}(x_{s-1} \notin T_0, x_s \in T_0)$$

Then since we move to every branch of the bottleneck with the same probability and the r branches $[0, n_{[k/2]}]$ are identical, L has the geometric distribution with parameter $\frac{1}{n_{[k/2]}}$. This leads to

$$(C2) \quad P(L \leq n_{[k/2]}) \leq \frac{1}{n_{[k/2]}}$$

Concerning (C3), the first assertion $\mathbb{E}_0(\tau_{\partial D}) \leq \mathbb{E}_{\partial D}(\tau_0)$ is trivially true with an equality since T_0 is by construction symmetric. It remains to determine conditions for (H)(b) and the modification of condition (H2) of Theorem 1.5. For (H)(b) we first notice that

$$(4.12) \quad \mathbb{E}[\theta_i] \lesssim \mathbb{E}[\lambda_{i,j}]$$

Since $(\theta_i)_i$ are distributed as the length of a random walk on r identical lines, with equal probability, of length z conditioning not to hit z and $(\lambda_{i,j})_{i,j}$ are iid random variables distributed as the length of a random walk on q identical lines, with equal probability, of length m conditioning not to hit m , the last inequality is true if and only if

$$(4.13) \quad m \geq z = n_{[\frac{k}{2}]}$$

which is true from (4.4). Because of (4.12), (H)(b) is reduced to

$$(4.14) \quad \mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L]\mathbb{E}[\lambda_{i,j}] \prec \frac{\gamma\sqrt{\text{Var}_c(\tau_0)}}{A}$$

We can bound the left hand side by

$$\mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L]\mathbb{E}[\lambda_{i,j}] \asymp m^2 + mn_{[\frac{k}{2}]} \lesssim m^2 = 6Nk$$

where above we made use of (4.13) and (4.4). So, for (4.14) to be true we need

$$6Nk \prec \frac{\gamma N \sqrt{k}}{A}$$

Since $A \leq \frac{1}{k^s}$ the last one is satisfied for every s such that

$$\frac{1}{2} - p < s$$

which is true for every $s \geq 1$. Finally for (H2) we need

$$\mathbb{E}_c(\tau_0) \leq \delta \sqrt{\text{Var}_{\partial D}(\tau_0)} \leq \frac{\mathbb{E}_c(\tau_0)}{A}$$

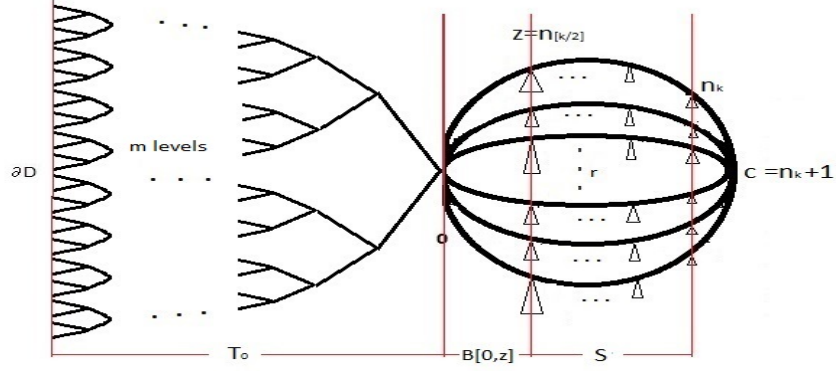


FIGURE 4. Paradigm 3

which is trivially true for any

$$(4.15) \quad t \leq s$$

4.3. paradigm 3. We will construct a graph based again on the example presented in [P-S]. For T_0 we will consider a binary tree of size M . The remaining part of the graph $B[0, z] \cup S$ will be the same with that of paradigm 2, as shown on figure 4.

We will determine r, γ, δ, m, M so that the conditions of Theorem 1.5 hold. As in the previous example we will start by placing conditions so that (H1) of Theorem 1.1 is not true. Since, in a binary tree

$$(4.16) \quad \sqrt{\text{Var}_{\partial D}(\tau_0)} \prec \mathbb{E}_{\partial D}(\tau_0)$$

it is sufficient to have

$$(4.17) \quad \mathbb{E}_{\partial D}(\tau_0) > \mathbb{E}_c(\tau_0) = 6Nk$$

Since from (4.2) we know that for large $M \asymp 2^m$

$$m^2 \lesssim \mathbb{E}_{\partial D}(\tau_0)$$

we can choose

$$(4.18) \quad m > \sqrt{Nk}$$

As in paradigm 2, since the part $B[0, z] \cup S$ is common in the two examples, condition (C2) and

$$\gamma \sqrt{\text{Var}_c(\tau_0)} \prec \mathbb{E}_c(\tau_0)$$

for $\gamma = k^p$ with $p < \frac{1}{2}$ are satisfied. Similarly,

$$(4.19) \quad t_S \prec 6Nk$$

and $A \leq k^{-s} < 1$ for some $s > 0$ if $\Phi = t_S \sum_{x \in S} \sum_{y \notin S} \mu_S(x) P(x, y) \leq k^{-2s} < 1$. To determine parameters for $\Phi < \frac{1}{k^{2s}}$ we compute

$$(4.20) \quad \sum_{x \in S} \sum_{y \notin S} \mu_S(x) P(x, y) = \frac{2}{(2+r)|S|}$$

where the size of S is

$$(4.21) \quad |S| \asymp rN \left(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j} \right)$$

Combining together (4.19), (4.20) and (4.21) we get

$$\Phi \asymp \frac{6Nk2 \left(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j} \right)^{-1}}{(2+r)rN} \leq \frac{6k \left(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j} \right)^{-1}}{r}$$

We can then choose

$$(4.22) \quad r = 6k^{2s+1} \left(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j} \right)^{-1}$$

Furthermore, we can request for s to be sufficiently large so that

$$k^s > \frac{m^3}{kN}$$

From the last one and the right hand side of (4.2) $\mathbb{E}_{\partial D}(\tau_0) < m^3$, we obtain

$$(4.23) \quad \frac{\mathbb{E}_c(\tau_0)}{A} > \mathbb{E}_{\partial D}(\tau_0)$$

Concerning (C1) we require (4.10)

$$|S| + |B[0, z]| \prec |T_0|$$

We have

$$|S| + |B[0, z]| \asymp rN \left(\sum_{j \in \{[k/2], \dots, k\}} \frac{1}{n_j} \right) = 6k^{2s+1}N$$

where above we used (4.22). Since the size of the binary tree T_0 is $|T_0| = M$ for (4.10) we need

$$6k^{2s+1}N \prec M \asymp e^m$$

Concerning (C3), both assertions are trivially true. For inequality (H)(b) we can compute

$$\mathbb{E}_{\partial D}(\tau_0) + \mathbb{E}[L]\mathbb{E}[\lambda_{i,j}] + \mathbb{E}[L]\mathbb{E}[\theta_i] \prec m^3$$

and

$$\frac{\gamma \sqrt{\text{Var}_c(\tau_0)}}{A} = k^{s+p+\frac{1}{2}}N$$

From the last two we obtain (H)(b) for s large enough such that

$$m^3 < k^{s+p+\frac{1}{2}} N$$

It remains to show the following two inequalities.

$$\mathbb{E}_{\partial D}(\tau_0) \prec \delta \sqrt{\text{Var}_{\partial D}(\tau_0)}$$

and

$$\mathbb{E}_c(\tau_0) \leq \delta \sqrt{\text{Var}_{\partial D}(\tau_0)} \leq \frac{\mathbb{E}_c(\tau_0)}{A}$$

However, since (4.16), (4.17) and (4.23) hold, for the last two inequalities to be true we need to choose δ in the following range

$$\frac{\mathbb{E}_{\partial D}(\tau_0)}{\sqrt{\text{Var}_{\partial D}(\tau_0)}} < \delta < \frac{\mathbb{E}_c(\tau_0)}{A \sqrt{\text{Var}_{\partial D}(\tau_0)}}$$

REFERENCES

- [L-P-W] DAVID A. LEVIN, YUVAL PERES AND ELISABETH L. WILMER, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009.
- [P-S] YUVAL PERES AND PERLA SOUSI *Total variation cutoff in a tree*. Annales de la Faculte des Sciences de Toulouse (to appear)